以 AIGC 为代表的人工智能在传媒领域的发展和应用

杨孔威
（新华社浙江分社，浙江 杭州 310007）

摘要：【目的】探讨人工智能给新闻传媒领域带来的变革和创新力量。【方法】分析新闻内容生成模式的变迁和 AIGC 爆发的原因及 AIGC 的实际应用。【结果】技术的发展促进了 AIGC 在新闻传媒领域的发展，AIGC 给新闻传媒领域带来了革命性的力量。【结论】传统媒体需要勇于创新，以更开放、积极的态度去拥抱以 AIGC 为代表的人工智能新技术。

关键词：人工智能；AIGC；深度学习；内容生成；神经网络　　中图分类号：G213　　文献标识码：A
文章编号：1671-0134（2023）05-076-05　　DOI：10.19483/j.cnki.11-4653/n.2023.05.017
本文著录格式：杨孔威 . 以 AIGC 为代表的人工智能在传媒领域的发展和应用 [J]. 中国传媒科技，2023（05）：76-80.

人工智能对于人们来说并不陌生。

早在 1997 年，超级计算机“深蓝”打败了国际象棋大师加里・卡斯帕罗夫，让大家惊叹于计算机的强大。2017 年 AlphaGo 以三比零战胜人类顶尖棋手柯洁，AlphaGo 超越了人类最强棋手的能力，带给人们极大的震撼。

自诞生以来，人工智能发展共经历了三次浪潮。1956 年科学界提出了人工智能的概念，人工智能之父艾伦・图灵提出了著名的“图灵测试”，受“图灵测试”的刺激，全球范围内出现了第一波人工智能的发展浪潮，但受当时技术的限制，人工智能仍处于科学探索阶段。20 世纪 80 年代，人工智能迎来了第二次发展浪潮，科学家们利用基于概率统计模型的新方法，通过分析语音、音素、音节之间的对应关系使得语音识别、机器翻译实现了新的突破，但这个阶段的数据量训练有限，无法进行有效扩展导致识别率较低，人工智能还不具备太多的实用价值。直到 2006 年，被称为“神经网络之父”的 Hinton 提出了深度学习技术，人工智能的第三次发展浪潮从此开启并迎来了跨越式的发展。

现阶段人工智能在计算机视觉、语音识别、自然语言处理、数据挖掘等领域都取得了突破，也不断地深入新的创新领域，呈现了深度学习、跨界融合、人机协同、群智开放、自主操控等新特征，对新闻生产模式和新闻编辑都产生了重大的影响。[1]

1. 新闻传媒领域生成模式的发展和变迁

在传统媒体时代，以报纸、电视为代表的媒体的生产模式是专业生产内容即 PGC（Professionally Generated Content）。进入互联网时代，一些知识背景和专业资质较高的个人或者团体按照几乎与传统媒体制作节目无异的方式制作内容，生产流程比较专业化，制作的内容精良、质量可控性较强。优酷、土豆是最早发力于 PGC 的视频网站，一些内容平台如微信公众号、门户网站也是属于 PGC 的生产模式。

互联网的普及，5G 技术的成熟促进了移动互联网的蓬勃发展，移动上网用户的不断增加带来了社交媒体的兴起，互动性的常态化使用户能以多种形式在网络上发布、展示和传播由用户个人创作的一些原创性的文字、图片或音视频内容。在信息大爆炸的时代，特别是伴随着以个性化为主要特点的 Web2.0 概念兴起和抖音、今日头条等 App 的火爆，使得人人都是内容生产者的 UGC（User Generated Content，用户生成内容）生产模式开始冲击传统媒体普遍采用的 PGC 内容生产模式。UGC 模式的出现使得社交媒体真正成为一种“所有人对所有人的传播。”[2]

早在人工智能的第二次发展浪潮时，科学家们就开始探索如何利用人工智能技术来生成各种类型的内容，例如新闻、音乐、诗歌等。这一阶段的人工智能还属于基于规则的、通过预先设定基础算法来生成内容的一种简单形式。随着 AI 技术的不断迭代，新型的内容生产方式 AIGC（AI Generated Content，人工智能生成内容）逐渐从概念走向了落地，并以超出人们预期的速度在各行业特别是传媒领域创造了新的格局和生态，深刻改变了行业的演进模式。

2. AIGC 爆发式发展的原因

2.1 内容生成的需求

在传播的实践中，内容生产永远处于整个传媒产

业链和价值链的中上游位置，掌握了内容优势地位的媒体，往往能够凭借优质内容在媒体竞争中立于不败之地。[3] 可以看到，在 PGC 模式下，内容生产依赖于专业化团队的制作，为了保障生成内容的质量，PGC 团队需要投入大量的人力、物力资源及技术成本，内容生产的高要求导致了 PGC 难以满足用户的数量和覆盖面需求。UGC 模式模糊了生产者与消费者之间的界限，用户既是内容生产者也是产品的受众和消费者，这种方式降低了用户内容生产和产品的准入门槛，满足了用户个性化和多样性需求，使得内容产品得以爆发式地增长，大大提高了媒体产业链的繁荣度，但由于 UGC 生产者专业化程度参差不齐，其内容质量不可避免地受到影响。UGC 时代以来，“新闻传播者”的角色不再只是专业新闻单位独有而可以是任何人，甚至有可能不是人，如此带来新闻信息良莠不齐。[4]

从用户角度而言，碎片时间的阅读习惯也让人们对阅读资源的需求飞速增长。在“个人媒体”急剧扩展的时代，用户通过智能终端从网络实时获取信息，也随时随地的传送和分享信息，信息发布的形式“千人千面”，且门槛低、渠道多，在带给受众更高的参与度和互动性的同时也造成了海量信息的淹没，参差不齐的信息海洋给用户增加了更多的搜索成本。相较于传统纸媒，智能终端更是带来传统纸媒单调图文信息所无法充分体现的视觉和听觉效果，人们需要优质的音视频内容带来立体的情感体验。种种原因使人们对信息内容的丰富性、多样性、专业性提出了更高的要求，而内容消费的强大需求与内容供给不足产生的矛盾依然突出，PGC、UGC 也逐渐难以匹配内容需求的极速扩展，AIGC 顺应内容的需求得到快速的发展。

2.2 技术的日益成熟

科技创新是先进生产力发展的主要动力，技术也是媒体发展变革的第一生产力。AIGC 的发展成熟离不开相关技术的进步。一般而言，AI 三大核心要素分别是数据、算力和算法。

2.2.1 海量的数据基础

海量数据是深度学习算法的基础。深度学习算法本质上是模拟了人类大脑的运行机制，而人类如果要获取相关的技能，就必须经过不断地实践，AI 也是如此。AI 算法的实现过程需要经过训练和推断两个阶段，训练阶段需要海量的数据输入，来训练出一个复杂的神经网络模型。经过完善的训练，这个神经网络模型可以快速高效地总结出规律，并从新的数据中“推断”出各种结论。得益于多层神经网络，参与训练的数据量越大，模型覆盖的场景越多，其精准度就越高，因此模型训练数据的丰富度在一定程度上决定了深度学习算法的优劣。

当今大数据的时代，网络、摄像头和传感器等设备无时无刻不在产生各种文字、语音、图片、视频等数据，据 IDC 的报告显示，随着互联网新技术的持续推进，2021 年全球数据量已经达到 82ZB，预计到 2025 年将达到 175 ZB[5]，这些数据都可成为 AI 算法充裕的训练“饲料”。伴随着大数据技术水平的不断提升，可用于标注的 AI 训练数据获得成本也得到降低，从而降低 AI 商业化的成本，也在一定程度上加速推进 AI 在各行业的普及应用。

2.2.2 算力能力的不断提升

算力是数字经济时代新的生产力，算力的大小与数据处理能力的强弱息息相关。算力源于芯片，作为算力的基础和核心硬件，芯片的发展程度决定着 AI 的进步程度。

早期，受限于芯片的处理能力，机器无法处理复杂的程序，人工智能只能完成简单的任务处理，计算机的性能严重制约了人工智能的发展。在过去的 20 年里，处理器性能以每年大约 55% 的速度提升，全世界的算力规模保持高速的增长态势，算力芯片的快速迭代使得本轮人工智能得以爆发。目前全球 AI 算力主要是以 GPU 芯片为主，随着硬件的发展和技术的不断更新，专用的 AI ASIC 芯片、定制化的 FPGA 芯片等有望成为新一轮 AI 算力发展的支撑，异构技术架构、专用计算架构、泛在协同计算架构等多样化的技术架构，以及计算技术与数学、物理等多学科交叉融合衍生出的量子计算、存算一体等计算技术的突破进展，将进一步推动着算力的加速提升。

2.2.3 算法的不断进步

20 世纪 40 年代，人们就提出了神经元数学模型的概念，80 年代，Hopfield 神经网络和 BT 训练算法的提出，进一步推动了神经网络的研究。早期 Hopfield 神经网络是通过抽象人脑结构和外界刺激响应机制，探索模拟人脑神经系统的学习、联想、记忆和模式识别等功能，从而进行逻辑操作的数学模型。[6] 但早期神经元网络只能解决单一的问题，复杂数据的处理受到限制，无法穷举复杂的情景，当网络规模和数据量积累到一定程度后，即使再投入更多的数据，准确率依然无法提升，实际应用价值有限。

2006 年，Hinton 等人提出了神经网络深度学习（Deep Learning）算法，将人类从复杂的算法归纳中

解放了出来，开启了机器学习的新领域和第三次人工智能的发展浪潮。与传统的机器学习算法比较，Deep Learning 算法的特点在于具有优异的特征学习能力，不需要人工去提取规则特征，机器能够实现特征的自动提取，将复杂的归纳算法留由机器去完成，同时，随着数据规模的增加，算法准确性也在不断提高，弥补了传统 BP 神经网络准确率不高的缺陷。相关数据显示，在引入深度学习之前，语音识别的准确率连续三年稳定在 76.4%，而 2018 年阿里巴巴达摩院语音识别团队利用深度学习技术推出的新一代语音识别模型 DFSMN 将全球语音识别准确率纪录提高至 96.04%。[7]

人工智能领域顶级专家吴恩达提到：发展人工智能就像用火箭发射卫星，需要强大的引擎和足够的燃料，算法模型就是其引擎，高性能的算力是打造引擎的工具，海量的数据就是引擎的燃料。“数据”“算法”“算力”是人工智能最重要的三大要素，它们之间相互促进和相互支撑，最终促成了人工智能技术的应用和价值创造。[8]

目前 AI 被广泛认为是人类历史上第四次工业革命，美国、中国等科技大国均将其提升至顶层战略高度，在技术的促进、政策的加持和资本的催化下，“弱人工智能”的应用已经快速渗透到各行业中，AIGC 也进入了新闻传媒领域。

3. AIGC 在新闻领域的应用

多模态的预训练模型给 AIGC 带来了多样化的内容生产能力，使其不仅能够拥有文字、图片、音频和视频等基本的内容生成模式，还可以对文字、图片、视频等相互之间进行跨模态的内容生产，给新闻传媒行业带来了一轮新的生产革命。

3.1 AIGC+ 文本生成

作为 AIGC 最早发展的技术，文本生成已经在新闻报道领域得到广泛的应用。早在 2014 年，《洛杉矶时报》就使用地震新闻自动生成系统播报了加州地震的第一条新闻报道，给新闻传媒行业带来了革新力量。近几年，随着人工智能技术的不断发展和成熟，国内外媒体都开始将其运用到新闻生产和传播的产业链中。美联社自 2018 年开始就用生成式 AI 工具 Wordsmith 自动生成体育新闻和财经报告，目前至少已经实现 50000 篇文章的自动化撰写。路透社在 2018 年启用了一款名为 Lynx Lnsight Service 的 AI 新闻撰稿工具，以帮助新闻记者分析数据、提出故事构想、自动生成有关金融市场和企业收益的报告。新华社于 2015 年就研发了一款名为“快笔小新”的新闻机器人，通过数据采集、数据加工、自动写稿、编辑签发从而实现提高新闻的制作效率。人工智能专家吴恩达表示：“AIGC 可以帮助人类创造更多高质量的内容，并且可以帮助人们更好地理解复杂的数据和信息。”国内外大型媒体机构都开始使用 AIGC 来提高新闻生产制作的速度和效率，AIGC 参与生产的内容也日益渗透到人们生活的方方面面，为读者提供了更加个性化的新闻体验。

3.2 AIGC+ 视频剪辑

随着 5G 技术的发展和智能终端的普及，受众“碎片化”内容消费习惯的形成使得短视频得以快速发展并成为当今各个内容消费领域的主流，但内容形式模式化和产品同质化也使得生产者面临着激烈的竞争和挑战。如何提高内容生产的效率，快速推出受众欢迎且高质量的内容成为各大媒体内容创作者的关注重点。AIGC 技术与短视频内容创作的结合带来的便利和优势成为问题的最佳解决方式，通过 AIGC 技术的应用，可以更好地提高视频的质量及创作效率，有效地区分目标受众，并且更加高效地将相关的作品推送出去。

AIGC 目前在视频领域的应用主要集中在视频内容属性的编辑和内容生产的自动剪辑等功能上。对于视频内容编辑，AIGC 可以实现自动画质修复、敏感人物识别、主题自动跟踪剪辑、画面特效、自动美颜等；对于视频自动剪辑，AIGC 可以基于视频中的画面、声音等多模态信息的特征进行解析，按照相应的语义限定进行检测，对满足条件的片段进行剪辑合成，从而实现智能提取、自动制作、全景直播拆条等功能。

早在 2017 年，新华社和新华智云立足新闻内容生产自动化场景，联合推出了名为“媒体大脑”的人工智能平台，利用 AIGC 技术帮助编辑快速锁定镜头，精彩片段快速拆条，生成内容一键快速发布到各大平台，简化了流程也为编辑节省时间实现了“快速传播”的功能。

中央广播电视总台在北京 2022 年冬奥会上采用 AI 智能自动化生产剪辑系统，利用海量赛事资源，对赛事关键时刻实现快速自动剪辑，规模化地自动生成并发布短视频内容，有效地节省人力成本，将编辑从繁琐的人工剪辑工作中解放出来。AIGC 的剪辑能力让中央广播电视总台在冬奥视频报道中实现了速度与质量的比拼优势。

AIGC 在体育媒体视频内容生产领域得到广泛应用是大势所趋，在大幅提升内容生产效率的同时还将向内容多元化方向进一步延伸，创建系统化、结构化的精品内容，满足受众对内容质量和数量的双重刚性需

求。

3.3 AIGC+ 数字人

AIGC 数字人在本质上也是建立了人与虚拟世界的链接和交互，同时解放了真人的劳动力。AI 数字人可以很好地根据真实人的外貌、动作、表情、声音等特点进行模拟以达到惟妙惟肖的程度，并且可以通过自然语言模型来模拟人类思维和行为特征。由于 AIGC 数字人是通过计算机创建的，它们不会生老病死，也不受时间和环境的影响，能够成为不眠不休的“劳模”。

通过自然语言处理、语音合成和语音识别技术的加持，在播音主持领域，AIGC 数字人能够像真人主播一样提供出色的播音主持工作，还能 24 小时在线，不仅能够根据场景节目创建不同的数字主持人，也可以模仿某个受用户喜爱的主持创建不同的分身，“扮演”不同栏目的主持人、新闻主播等角色，分别讲解科技、文化、历史、地理、美食等不同领域的知识，“他们”风度翩翩、学富五车，不仅拥有了全面、广泛的知识，如果接入对话模型（ChatGPT）数字人还可以实现面对面的交流，回答观众提出的各类问题，让人机交互更加真实和惟妙惟肖。人机协同不仅仅表现在信息处理流程方面，更多表现的是在身心一体的深度情感沟通层面上。[9]

新华社早在 2020 年就联合搜狗公司推出了全球首个 3D AI 合成主播“新小微”，其采用超写实 3D 数字人建模、实时面部动作生成及驱动、多模态识别及生成、迁移学习等多项人工智能前沿技术，根据输入的文本机器可以自动生成相似度极高的数字人视频内容，同时在播报过程中能够根据语义产生相对应的面部表情和肢体语言。[10]

新华智云 2019 年便开始试水数字人，开创了新闻领域实时音频与 AI 真人形象合成的先河，基于深度学习模型、动作模拟、情感模拟等科技，AI 通过采集几分钟的真人视频进行几小时的训练后，生成形象逼真、表情到位、口型匹配的数字人。在 2020 年的地方两会报道中，就有 7 省的两会报道使用了新华智云的虚拟主播。

2023 年全国两会期间，百度运用了可交互式超写实数字人与 AIGC 技术，将数字人与人工智能生成内容相结合，以科技感十足的人机交互式对话方式，向公众在线解读最高人民法院工作报告。

随着技术的成熟，依托 AI 技术驱动的数字人将成为未来数字人市场的主流。越来越接近真人外形的数字人，将给各行各业特别是新闻传媒行业的受众提供更亲切、自然、高效的服务体验。AI 浪潮之下，AIGC 数字人也将成为人类链接数字世界的纽带，加速元宇宙等产业的持续演进和多元化发展，丰富数字化应用场景并推动数字经济产业的发展。IDC 在《中国 AI 数字人市场现状与机会分析，2022》报告中预计，到 2026 年中国 AI 数字人市场规模将达到 102.4 亿元。[11]

3.4 AIGC+ 声音应用

在新闻媒体的传播过程中，声音以其独特的吸引力，成为不可或缺的“元素”，专业的配音传达了新闻叙述者的情感，提升了观众的共情和体验感。然而音频制作本身也存在一些难点，例如传统节目中由创作者制作的音频不仅创作形式单一，而且对配音员要求高，耗时长，配音成本高。随着人工智能技术的发展，AI 语音识别、语音合成等技术逐步在新闻媒体领域得到应用。早期的语音生成系统因为缺乏表达逻辑推理和因果关系的能力，缺乏连续语音的韵律，机械感重让人感觉单调不真实；近年来，随着数字信号处理技术的飞速发展，语音合成技术也取得了相当大的进步，高度拟人、流畅自然的语音合成服务、语音播报、仿真真人配音也在新闻媒体中得到广泛应用并提高了音频内容的用户体验感。

AIGC 也开始在语音克隆、生成虚拟人的定制声音等领域得到应用，并且交互性、实时性进一步增强。生成的音频内容有情绪、有温度，或深层雄厚，或俏皮可爱，或铿锵有力，或绵柔动人。

科大讯飞 2023 年新发布了一条关于节气《雨水》的视频，深沉厚重的男低音自带质感，这段配音由科大讯飞的 SMART-TTS 系统合成，其语气的变化、语句的停顿、声音的细腻程度都和真人所差无几。喜马拉雅利用 AIGC 理解文本的语境，选择合适的音色，根据文本的情绪随时转换声音，创造了多情感、多风格的声音模型运用在新闻、小说、财经等不同类型内容的制作中。

2022 年 11 月 1 日，新华社利用 AIGC 能力——人工智能演唱及智能视频创作，发布了数字记者、全球首位数字航天员小诤的单曲 MV《升》，此次发布的 AI MV《升》由新华社媒体融合生产技术与系统国家重点实验室联合腾讯音乐娱乐集团出品，歌曲演唱体现了高度拟人化的合成语音技术，生成的语音甜美且深富情感。[12]

AIGC 与智能语音技术的深度结合，以及应用场景在新闻媒体行业中的创新落地，有望进一步推动智能语音产业市场的发展。

3.5 AIGC+ 视频生成

AI 生成配音可以说是当下各种媒体行业特别是短视频内容创作者最为依赖的工具，但随着 AIGC 的跨越式发展，传统的配音已经满足不了 AIGC 的“大放异彩”之心了。从技术的发展可以看出，在语言大模型、图像大模型之后，基于深度学习的生成扩散模型和多模态预训练大模型已经成为新的趋势，视频生成就是其中具有代表性的应用领域。

2022 年 Meta 公司推出了名为 Make-A-Video 的文本生成模型，实现了通过文本直接“书写”视频的能力，其 AI 模型还能实现将两张静态图像转成视频，或者根据两张图片生成一段连续视频，以及基于一段原视频生成新视频的能力，生成的视频具有丰富的审美和艺术风格。同样谷歌公司也推出名为 Phenaki 的 AI 模型，只需向其提供一段剧本提示词，这个文本转视频（Text-to-Video）模型就能生成一段长达两分钟且颇具故事性的视频内容。

国内 AI 巨头百度也不甘落后，早在 2019 年就已开始积累 AI 预训练模型技术，并把文心大模型的能力也运用到了智能视频合成平台 VidPress 中，实现图文自动转换视频的功能，即把文字分析与摘要、媒体素材收集、素材智能化处理、音视频对齐，以及视频剪辑这 5 个步骤实现自动化，其中使用了大量的自然语言处理技术对图文内容进行分析，这些都是基于文心大规模训练实现的。

2022 年，新华社与百度文心一格联合推出了一支 AIGC 视频《AI 描绘天宫盛宴》，视频画面全部由 AI 自动生成，通过一幅幅具备东方意象的恢宏绚丽画作，回顾了 30 年来中国载人航天的辉煌历程。

在传媒、文化娱乐、教育等诸多领域，生产创作视频内容都是当下吸引受众眼球最重要的方式，越来越多的人通过制作创意短视频内容进行推广从而实现商业化变现，这些都将促使各行业对基于 AIGC 的视频生成内容有强烈的需求，在视频成为信息主要表达载体的当下，无论在新闻节目、短视频、直播，还是影视等领域，AI 视频内容生成都将成为 AIGC 的主要发展方向。当视频生成质量提高到专业水平时，AIGC 也将打开更多内容创作的想象空间，更大规模地服务受众，满足现代人的生活需求。

结语

伴随人工智能等新技术的不断完善，媒体已经进入一个全新的智媒时代，媒体智能化正在重塑传媒业的原有生态。[13] 正如比尔盖茨所说，AIGC 这种变革性技术，其影响力不亚于 PC 或互联网诞生。AIGC 在科技界掀起的惊涛骇浪正在向各个领域延伸，给人们的工作、生活模式带来巨大变化。它的出现给包括新闻媒体在内的各行各业带来了新的挑战和机遇，也促使人们必须以更加开放、积极的态度去适应技术的不断创新和发展，不断提升自身的创新能力和核心竞争力，如此才能在这股科技的大浪淘沙中立于优势地位。

参考文献

[1] 任瑞娟，王保超，赵雅倩 . 演进与动向：人工智能在传媒领域的应用 [J]. 新闻与传播评论，2021（2）：26-35.

[2] 王艺焜 . 社交媒体的传统媒体化：从内容生产角度的调查与分析 [EB/OL].http：//media.people.com.cn/n1/2017/0110/c409682-29013077.html.2017-01-10/2023-03-25.

[3] 邹海涛 . 对融媒体时代“内容为王”的思考 [J]. 今传媒，2017（11）：150-151.

[4] 樊虹 赵小波 .“人工智能 + 传媒”：重构与再造新闻生产流程 [J]. 军事记者，2020（8）：22.

[5] IDC Innovator：智能文档处理（IDP），2022[EB/OL]. https：//www.sohu.com/a/630711212_121124366.2023-01-16/2023-03-25.

[6] 人工神经网络简介 [EB/OL].https：//blog.csdn.net/fengbingchun/article/details/50274471.2015-12-12/2023-03-25.

[7] 阿里巴巴推出了新一代语音识别模型 DFSMN 已开源 [EB/OL].https：//www.sohu.com/a/235054085_223764.2018-6-11/2023-03-26.

[8] 孙仕 . 基于人工智能技术治理网络暴力的探析 [J]. 中国传媒科技，2023（1）：64-68.

[9] 赵馨 . 情意表达：AI 合成新闻直播的身体传播刍议 [J]. 中国传媒科技，2022（4）：65-67.

[10] 邬金夫 .“她”来了！全球首位 3D 版 AI 合成主播精彩亮相 [EB/OL].http://www.xinhuanet.com/politics/2020lh/2020-05/20/c_1126011533.htm2020-5-21/2023-03-23.

[11] IDC 报告：2026 年中国 AI 数字人市场规模将达 102.4 亿元 [EB/OL].http：//isuike.com/archives/15559.2022-06-28/2023-03-25.

[12]AI 演唱、AI 创作画画，新华社数字航天员发布单曲 [EB/OL].https：//baijiahao.baidu.com/s?id=1748275441565636192&wfr=spider&for=pc.2022-11-02/2023-03-26.

[13] 彭兰 . 万物皆媒——新一轮技术驱动的泛媒化趋势 [J]. 编辑之友，2016（3）：5-10.

作者简介：杨孔威（1977-），男，浙江杭州，硕士研究生，工程师，研究方向为媒体技术。

（责任编辑：张晓婧）